I. Narrative Description

This third year of the 5-year renewal of the SUMEX resource grant has been an active year not only for the SUMEX staff, but for the SUMEX-AIM community involved in developing expert systems. Successes in developing such systems, many of them stemming from projects in the SUMEX-AIM community, continue to stimulate strong and growing interest in AI research on many educational, governmental, and industrial fronts.

This is an annual report for the Stanford University Medical EXperimental computer resource for applications of Artificial Intelligence in Medicine (SUMEX-AIM). It covers the period between May 1, 1983 and April 30, 1984.

This third year of the 5-year renewal of the SUMEX resource grant has been an active year not only for the SUMEX staff, but for the SUMEX-AIM community involved in developing expert systems. Successes in developing such systems, many of them stemming from projects in the SUMEX-AIM community, continue to stimulate strong and growing interest in AI research on many educational, governmental, and industrial fronts.

In addition, this past year has seen concurrent development of SUMEX-AIM as a distributed scientific resource. Our approved project goals focus principally on the merging of state-of-the-art community research in biomedical AI applications with new computing tools and on the challenges they will bring to the SUMEX-AIM community and resource. The SUMEX staff continues to exploit these advances in professional workstations and communication technology, while at the same time maintaining our high standards for a computing resource.

This third year also saw the initiation of a number of SUMEX-AIM pilot projects. These pilot projects provide new activities and research directions for the community to replace existing projects which have matured and moved off the SUMEX-AIM resource.

The earlier phases of the SUMEX-AIM resource were characterized by the building of a national community of biomedical AI collaborators around a central resource located at Stanford University. Beginning with 5 projects in 1973, the AIM community grew to 11 major projects at our renewal in 1978. This past year saw the completion of two long term and successful projects on SUMEX-AIM; DENDRAL and PUFF/VM. There currently are 13 fully-authorized projects plus seven pilot efforts.

Many of the computer programs under development by these groups are maturing into tools increasingly useful to the respective research or clinical communities. We continue to seek out new AI applications in our community of biomedical and computer scientists who interact through electronic media. The SUMEX-AIM community is beginning to evolve as a highly distributed resource, with the SUMEX staff and computer facility serving as the backbone to electronic communication and systems support. The community is becoming more and more involved in personal computers and professional workstations, and more heavily dependent on network communication facilities for interactions, collaborations, and sharing.

The following sections cover the activities of the SUMEX-AIM resource this past year, including brief summaries of the our objectives, a characterization of biomedical AI research, resource organization and operating procedures, recent core progress in system development and basic AI research, and progress in the collaborative projects.

I.A. Summary of Research Progress

I.A.1. Overview of Objectives and Rationale

SUMEX-AIM ("SUMEX") is a national computer resource with a dual mission: 1) promoting applications of computer science research in artificial intelligence (AI) to biological and medical problems, and 2) demonstrating computer resource sharing within a national community of health research projects. The central SUMEX-AIM facility is located physically in the Stanford University Medical School and serves as a nucleus for a community of medical AI projects at universities around the country. SUMEX provides computing facilities tuned to the needs of AI research and communication tools to facilitate remote access, inter- and intra-group contacts, and the demonstration of developing computer programs to biomedical research collaborators.

I.A.1.1. What is Artificial Intelligence

The subfield of computer science known as Artificial Intelligence, or AI, deals with symbolic reasoning using large amounts of heuristic knowledge. Many of the world's difficult problems are symbolic, such as troubleshooting electronic or mechanical equipment, medical diagnosis and therapy planning, and configuring elemental parts into a whole system. For these kinds of problems, AI offers new opportunities for developing computer-based solutions.

In addition to using symbolic representations of knowledge, AI also uses heuristic methods for processing information. Heuristics are rules of thumb, judgmental rules that aid in finding *plausible* solutions. AI is distinguished from other areas of computing in its attention to both symbolic (non-numeric) information and heuristic (non-algorithmic) methods for solving problems.

Placing AI in Computer Science

The major focus of AI is understanding intelligence through construction (or programming) of machines that behave intelligently. That is a grand goal. In the short-term, AI research focuses on non-numerical problem solving in order to build experience with problem solving methods, techniques for representing various kinds of knowledge, interfaces with users, and numerous other issues.

One of the distinguishing features of problems for which AI methods have been developed is that the problems are not well-structured. That is, one does not already know in advance (from the problem description alone) what the best method is for solving the problem. In short, there are no algorithms. Broadly speaking, AI substitutes exploratory search for precise, algorithmic solution methods.

Expert Systems and Applications

The national SUMEX-AIM resource is an outgrowth of a long, interdisciplinary line of artificial intelligence research at Stanford and elsewhere concerned with the development of concepts and techniques for building "expert systems" [1]. An "expert system" is an intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solution. For some fields of work, the knowledge necessary to perform at such a level,

plus the inference procedures used, can be thought of as a model of the expertise of the expert practitioners of that field.

Two important features that distinguish expert systems from conventional programs are flexibility and understandability. Expert systems are *flexible* in the sense that they can be changed and extended easily, and they are *understandable* in the sense that they can explain the contents of their own knowledge bases and their own lines of reasoning [10]. These features are especially important in medicine, where knowledge is changing rapidly and where practitioners have to understand the reasons for a program's decisions because they have to accept responsibility for following (or not following) those decisions.

The application areas range from medicine to electronics, from machinery to software. The problems range from diagnosis and troubleshooting (analysis) problems to planning and configuration (synthesis) problems. Knowledge bases for expert systems are built iteratively -- usually through long interactions over many months between a human specialist who understands the details of the domain and a knowledge engineer who understands the programming details of the system.

The knowledge of an expert system consists of facts and heuristics. The "facts" constitute a body of information that is widely shared, publicly available, and generally agreed upon by experts in a field. The "heuristics" are the mostly-private, little-discussed rules of good judgment (rules of plausible reasoning, rules of good guessing) that characterize expert-level decision making in the field. The performance level of an expert system is primarily a function of the size and quality of the knowledge base that it possesses. One of the key ideas in maintaining flexibility and understandability is the clean separation of elements of the knowledge base from elements of the program that interpret the knowledge base.

The major issues in building expert systems, at the moment, are:

- selecting an appropriate problem (in terms of size, difficulty, importance, decomposability, risk)
- selecting a representation and control structure (or framework system that supplies both),
- settling on an appropriate vocabulary and conceptualization for the problem,
- finding an available expert,
- transferring the expert's knowledge into the program (knowledge engineering),
- refining the knowledge base with feedback from test cases,
- packaging the system in a form that is acceptable to end-users,
- validating the quality of the program's advice.

One of the best known expert systems is MYCIN [3], a program in which the separation of knowledge (of medicine) from the rest of the program was carefully engineered. (The abstracted case of an arbitrary knowledge base and a framework interpreter, plus auxiliary programs, was achieved in the EMYCIN system [16], to which knowledge of other domains can be added to build a diagnostic system in those domains.)

Currently authorized projects in the SUMEX community are concerned in some way with the application of AI to biomedical research[*]. The tangible objective of this approach is the development of computer programs that will be more general and effective

consultative tools for the clinician and medical scientist. There already have been promising results in areas such as chemical structure elucidation and synthesis, diagnostic consultation, molecular biology, and modeling of psychological processes.

Needless to say, much is yet to be learned in the process of fashioning a coherent scientific discipline out of the assemblage of personal intuitions, mathematical procedures, and emerging theoretical structure comprising artificial intelligence research. State-of-the-art programs are far more narrowly-specialized and inflexible than the corresponding aspects of human intelligence they emulate; however, in special domains they may be of comparable or greater power, e.g., in the solution of formal problems in organic chemistry.

I.A.1.2. Impact of AI in Biomedicine

There is a certain inevitability to the field of Artificial Intelligence and its applications, in particular, to medicine and biosciences. The cost of computers will continue to fall drastically during the coming two decades. As it does, many more of the practitioners of the world's professions will be persuaded to turn to economical automatic information processing for assistance in managing the increasing complexity of their daily tasks. They will find, from most of computer science, help only for those problems that have a mathematical or statistical core, or are of a routine data-processing nature. But such problems will be relatively rare, except in engineering and physical science. In medicine, biology, management, indeed in most of the world's work, the daily tasks are those requiring symbolic reasoning with detailed professional knowledge. The computers that will act as *intelligent assistants* for these professionals must be endowed with symbolic reasoning capabilities and knowledge.

The growth in medical knowledge has far surpassed the ability of a single practitioner to master it all, and the computer's superior information processing capacity thereby offers a natural appeal. Furthermore, the reasoning processes of medical experts are poorly understood; attempts to model expert decision-making necessarily require a degree of introspection and a structured experimentation that may, in turn, improve the quality of the physician's own clinical decisions, making them more reproducible and defensible. New insights that result may also allow us more adequately to teach medical students and house staff the techniques for reaching good decisions, rather than merely to offer a collection of facts which they must independently learn to utilize coherently.

The knowledge that must be used is a combination of factual knowledge and heuristic knowledge. The latter is especially hard to obtain and represent since the experts providing it are mostly unaware of the heuristic knowledge they are using. Medical and scientific communities currently face many widely-recognized problems relating to the rapid accumulation of knowledge, for example:

- codifying theoretical and heuristic knowledge

- effectively using the wealth of information implicitly available from textbooks, journal articles and other practitioners

- disseminating that knowledge beyond the intellectual centers where it is collected

- customizing the presentation of that knowledge to individual practitioners as well as customizing the application of the information to individual cases

We believe that computers are an inevitable technology for helping to overcome

these problems.   While recognizing the value of mathematical modeling, statistical classification, decision theory and other techniques, we believe that effective use of such methods depends on using them in conjunction with less formal knowledge, including contextual and strategic knowledge.

Artificial intelligence offers advantages for representing and using information that will allow physicians and scientists to use computers as intelligent assistants. In this way we envision a significant extension to the decision-making powers of specific practitioners without reducing the importance of those individuals in that process.

Knowledge is power, in the profession and in the intelligent agent. As we proceed to model expertise in medicine and its related sciences, we find that the power of our programs derives mainly from the knowledge that we are able to obtain from our collaborating practitioners, not from the sophistication of the inference processes we observe them using.   Crucially, the knowledge that gives power is not merely the knowledge of the textbook, the lecture and the journal, but the knowledge of *good practice*--the experiential knowledge of *good judgment* and *good guessing*, the knowledge of the practitioner's art that is often used in lieu of facts and rigor.   This heuristic knowledge is mostly private, even in the very public practice of science. It is almost never taught explicitly, is almost never discussed and critiqued among peers, and most often is not even in the moment-by-moment awareness of the practitioner.

Perhaps the the most expansive view of the significance of the work of the SUMEX-AIM community is that a methodology is emerging for the systematic explication, testing, dissemination, and teaching of the heuristic knowledge of medical practice and scientific performance. It may be less important that computer programs can be organized to use this knowledge than that the knowledge itself can be organized for the use of the human practitioners of today and tomorrow.

Evidence of the impact of SUMEX-AIM in promoting ideas such as these, and developing the pertinent specific techniques, has been the explosion of interest in medical artificial intelligence and the specific research efforts of the SUMEX community.   As SUMEX has entered its second decade, we have found that the small community of researchers that characterized the AIM field in the early 1970's has now grown to a large, accomplished, and respected research community.   The American Association for Artificial Intelligence (AAAI), the principal scientific membership organization for the AI field, has 4000 members, over 1000 of whom are members of the medical special interest group known as the AAAI-M.   This subgroup was founded by members of the SUMEX-AIM community who were active in AAAI and is the only active subgroup in the Association. The organization distributes semiannual newsletters on medical AI and provides a focus for co-sponsoring relevant medical computing meetings with other societies (such as the American Association for Medical Systems and Informatics -- AAMSI). Medical AI papers are prominently featured at both medical computing and artificial intelligence meetings, and artificial intelligence is now routinely featured as a specific subtopic for specialized sessions at medical computing and other medical professional meetings.   For example, members of the AIM community have represented the field to physicians at the American College of Pathology and American College of Physicians meetings for the last several years.   A mere decade ago, the words "artificial intelligence" were never uttered at such conferences.   The growing interest and recognition are largely due to the activities of the SUMEX-AIM community.

Another indication of the growing impact of the SUMEX-AIM community is its effect on medical education.   For reasons such as those outlined above, there is an increasing recognition of the need for a revolution in the way medicine is taught and

medical students organize and access information. Computing technology is routinely cited as part of this revolution, and artificial intelligence (and SUMEX-AIM research) generally figures prominently in such discussions. Such diverse organizations as the National Library of Medicine, the American College of Physicians, the Association of American Medical Colleges, and the Medical Library Association have all called for sweeping changes in medical education, increased educational use of computing technology, enhanced research in medical computer science, and career development for people working at the interface between medicine and computing; reports of all four organizations have specifically cited the role of artificial intelligence techniques in future medical practice and have used SUMEX-AIM programs as examples of where the technology is gradually heading.

In summary, the logic which mandates that artificial intelligence play a key role in enhancing knowledge management and access for biomedicine -- a logic in which we have long believed -- has gradually become evident to much of the biomedical community. We are encouraged by this increased recognition, but realistic about the significant research challenges that remain. Our goals are accordingly both scientific and educational. We continue to pursue the research objectives that have always guided SUMEX-AIM, but must also undertake educational efforts designed to inform the biomedical community of our results while cautioning it about the challenges remaining.

I.A.2. Details of Technical Progress

I.A.2.1. Facility Management and Operation

The following material covers the SUMEX-AIM resource activities over the past year in greater detail. Individual sections cover progress in ;

- Facility Management and Operation

- Timesharing Systems

- Professional Workstations

- Networking and Communications

These sections outline accomplishments in the context of the resource staff and resource management. Details of the progress and plans for our external collaborative projects are presented in Section II beginning on page 69.

I.A.2.2. Facility Management and Operation

SUMEX-AIM continues to manage and operate it's computing resources in a effective and efficient manner conducive to providing a reliable and robust computing environment.

While the previous year (Year 10) involved a major move from the KI10 Tenex system to a new DECsystem 2060, this year saw more emphasis to our gradual move to distributed processing, while continuing to improve our excellent timesharing environment on the 2060. This development is covered in full in section I.A.2.2 starting on page 15.

Our continued movement to professional workstations has taken on several forms. We have continued to acquire Lisp machines for use by the SUMEX community while at the same time investigating the use of remote virtual graphics and new lower cost workstations such as the Apple Macintosh, Sun workstations, and others that are appearing on the market. The development of professional workstations is covered in more detail in section I.A.2.3 starting on page 21.

SUMEX continues to expend a great deal of effort in the support and development of our networking and communications facilities. Key to our ability to provide the maximum computing power available to the greatest number of users is a mechanism for making it irrelevant where that user is physically located. By having a robust networking and communications environment, we are able to extend our facility to any user or group of users, thereby making available to them the power and convenience of SUMEX. Further information on the progress made in networking and communications can be found in section I.A.2.4 starting on page 23.

In the area of facility management and operation, several noteworthy events occurred over the past year which will be explained in more detail here.

E. A. Feigenbaum

SUMEX/HPP Welch Road Computing Facility

A major development this past year at SUMEX was the move of the Heuristic Programming Project to their new location at 701 Welch Road, adjacent to the Stanford Medical Center. Since this group is a major user of the SUMEX-AIM resource and the focus for most of the core AI research, a good deal of effort was expended to provide a robust computing environment at their new location. This development involved several stages and levels of technical development, ranging from construction of the machine room, new cable and wiring installation, procurement and setup of networking hardware, to major new developments in the networking software and a twisted pair Ethernet communication link between this site and the main SUMEX Computer room. All of the hardware and facilities purchases were funded from sources other than SUMEX.

We setup the general communications capabilities for the two buildings occupied by the HPP. This involved wiring up local terminals, installing local Ethernets (both 3 and 10 megabit capability), and acquiring and installing networking hardware such as terminal interface processors (TIP) and gateways, as well as extending the current SUMEX TIP and GATEWAY software to handle both 3 and 10 megabit network traffic.

But the most important and most interesting development in this process was the "twisted pair' ethernet developed by the SUMEX engineering staff to allow high speed reliable communications between this Welch Road facility and the SUMEX machine room. Further information on this new ethernet can be found in Section I.A.2.4 on page 23.

HPP researchers are routinely using this link to communicate with SUMEX and the central university network. In addition, various Lisp machines and printers located in the HPP facility and connected to a local network are able to communicate with the university network.

The end result is that we have successfully been able to extend the SUMEX computing environment to a remote site, providing a high speed link to the facilities of SUMEX while also allowing for local distributed processing. We see this experience has being most valuable in the future as we move further into a distributed environment, while still needing the sharing of resources and communication links provided by large timesharing systems and local area networks.

Digital Equipment Corporation stops development of 36-bit product line

Digital Equipment Corporation, a long time supplier of high speed 36-bit timesharing computers to the Artificial Intelligence community, announced that it was stopping all development of future 36-bit products, and instead starting a program to provide a migration path to its line of VAX minicomputers.

Many DEC 20 customers had been anticipating a new yet unannounced machine from DEC code named the 'Jupiter', which had been reported to be a order of magnitude faster than the current KL10 processor used in DEC20's and DEC10's. However, DEC's announcement means this effort has stopped, and we can expect no more 36-bit products from Digital Equipment Corporation.

The effect of this announcement to the AI community is disappointing, although not totally unexpected. The DECsystem20 has been the predominant timesharing machine used to support Artificial Intelligence based research, but yet researchers have been in need of more processing power and larger address spaces for quite a few years. DEC has clearly decided to devote their resources to VAX development. For those in need of greater 36-bit processing power or address space, you must now look to newer less

experienced companies such as Foonley and Systems Concepts for follow on 36-bit products which those firms are preparing.

The impact of this decision on the SUMEX-AIM community must be examined in conjunction with our development of AI systems on personal Lisp machines. We have outlined very clearly our plans to move to distributed Lisp-based workstations for AI Research, and this is clearly where we see the AI computing market heading. These machines offer much better cost/performance ratios than timesharing machines, high resolution bit-mapped screens, and powerful Lisp programming environments for the development and eventual dissemination of AI based systems. However, this is not too say we still do not see a role for the large timesharing machine in our environment. We still believe in the use of a large central mainframe computer as the anchor for a large community of users. The mainframe also functions as a central facility for communication and collaboration, and provides fast Lisp cycles for program development where the application is not in need of a specialized workstation.

Other SUMEX Computing Facilities

SUMEX continues to support other mainframe computers, file servers, professional workstations, and assorted printers and terminals for use by the SUMEX-AIM community.

1. The SUMEX-AIM File server, based on a VAX 11/750 computer, continues to serve the needs of the workstation users within SUMEX-AIM. The use of SAFE by users of our 2060 is minimal. We plan to extend the use of SAFE in the future by providing more convenient access by 2060 users than is currently available.

2. The VAX 11/780 computer system, originally purchased with DARPA funds and previously located in Margaret Jacks Hall on campus, has become a SUMEX-AIM resource this past year. The system was moved to a new location on the Stanford campus which provides a better environment for a computer of this size. This VAX is now shared between the Computer Science Department and the SUMEX-AIM community.

3. SUMEX continues to support a wide range of professional workstations from such vendors as Xerox, Symbolics, and Hewlett Packard for the development and testing of AI applications. Additional work has been started to explore the use of the Apple Macintosh and Apple Lisa within SUMEX. More information on these developments can be found in section I.A.2.3.
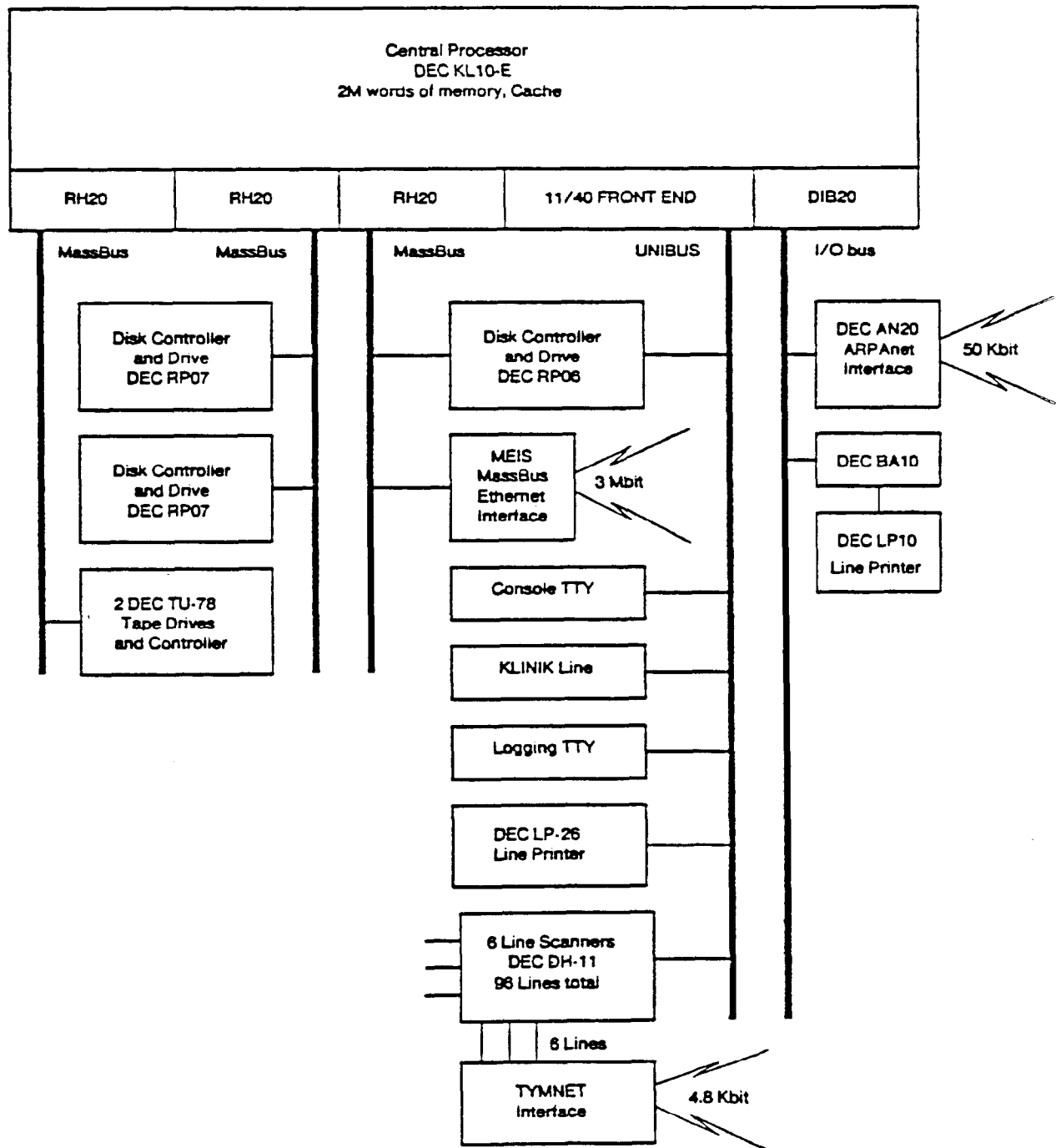
Figure 1: Current SUMEX-AIM Decsystem 2060 Computer Configuration
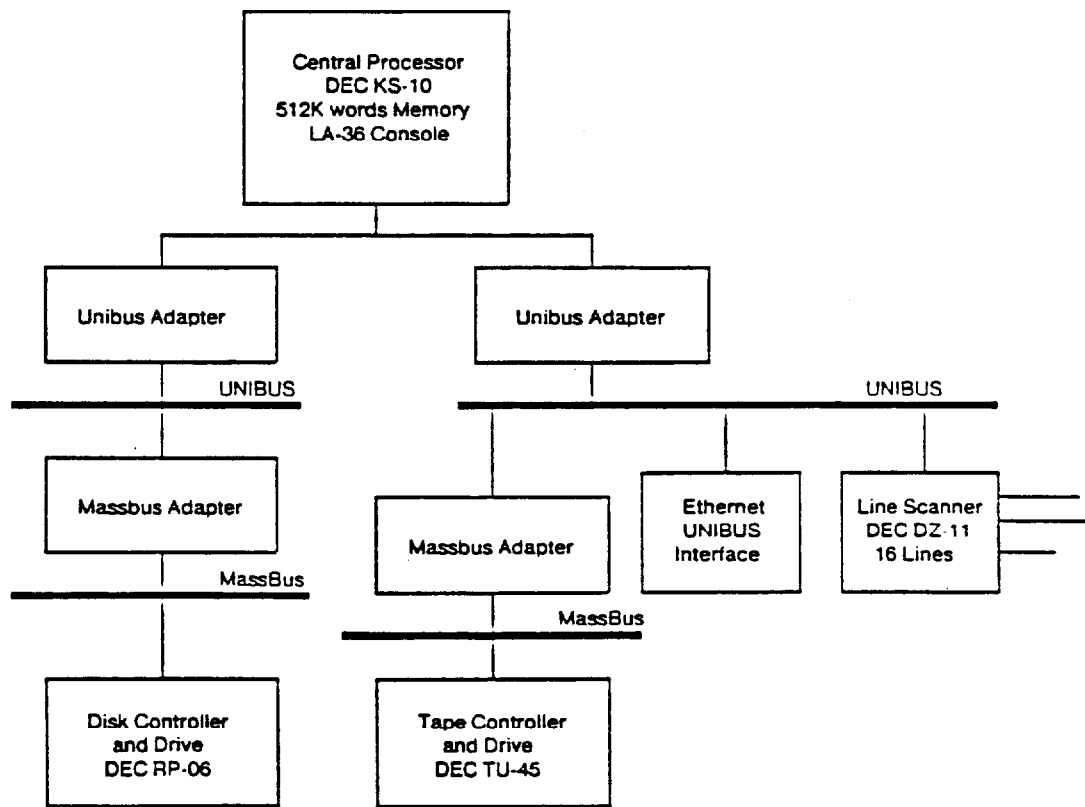
Figure 2: Current SUMEX-AIM 2020 Computer Configuration

Central Processor
DEC VAX 11/780
4 Mbytes Memory
Floating Pt Unit
LA-36 Console

MassBus   UNIBUS

Disk Controller
and Drive
DEC RP-06

Line Scanner
DEC DZ-11
16 Lines

Disk Controller
and Drive
DEC RP-07

Ethernet
UNIBUS
Interface

Tape Controller
and Drive
DEC TU-77

Chaosnet
UNIBUS
Interface
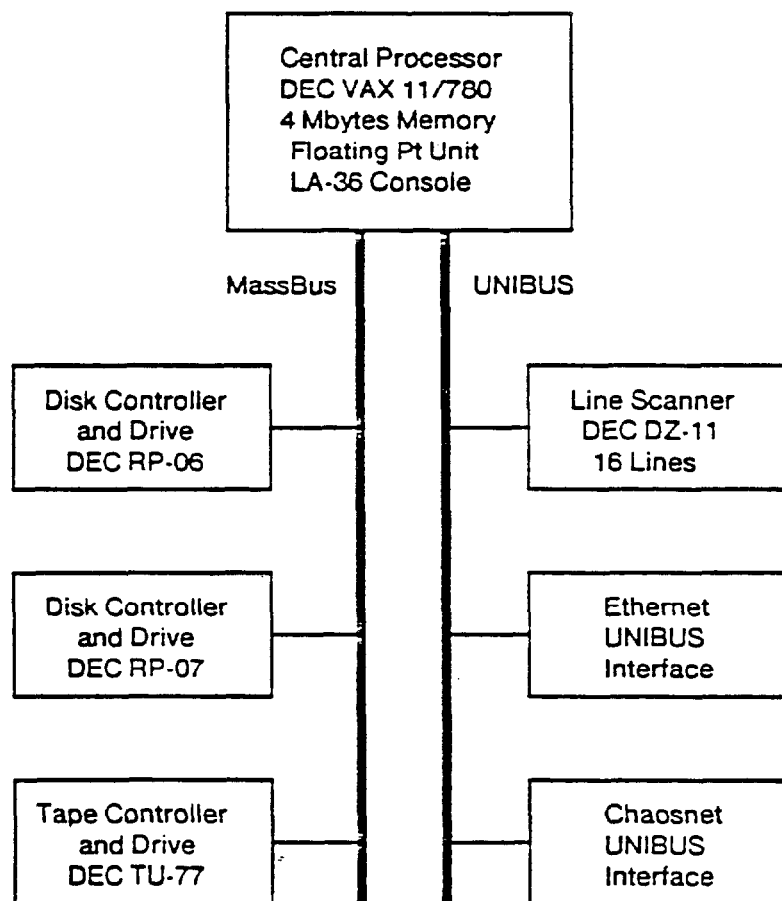
Figure 3: Current Shared VAX Computer Configuration

Margaret Jacks
Heuristic Prog Proj

Electrical
Engineering

Pine Hall

CSLI

10 Mbit ethernet

SCORE 2060
Ether TIP
Dover Printer
Other CSD Equip

Sumex VAX
CSLI 2060
Dandelions
NS file server

G

G

R

G

Oncocin Offices

Ether TIP
Dandelions
HP 9836's

R

SUMEX Offices

Medical Center
SUMEX Mach Rm

2060
2020
VAX - File Server
VAX - Devel
Ether TIP
PDP-11 - Devel

Xerox Alto
Dolphin
Dandelions
SUN - Devel
Ether TIP
Imprint-10 Printer

R   Repeater

G   Gateway

10 Mbit ethernet

G

1.5 Mbit
phone line

Whelan Building (Welch Road)
Heuristic Prog Proj

10 Mbit ethernet

Dandelions
LM-3600's
Ether TIP

G

Dolphins
Dorado
Raven Printer

3 Mbit ethernet

Figure 4:   SUMEX-AIM Ethernet Configuration

```
                            ┌─────────────────────┐
                            │ Central Processor   │        ┌──────────────┐
                            │ DEC VAX 11/750      │────────│ Console TTY  │
                            │ 2 Mbytes Memory ┌───┤        └──────────────┘
                            │                 │TU 58│
                            └─────────────────┴───┘
                                MassBus      UNIBUS

  ┌──────────────┐                                    ┌──────────────┐
  │ CDC 256 Mbyte│   ┌──────────────┐                 │   Kennedy    │
  │ Removable-   │   │ Fujitsu Eagle│                 │ 800/1600 bpi │
  │   Media      │   │ 414 Mbyte    │                 │  Tape Drive  │
  │ Disk Drive   │   │ Disk Drive   │                 │  & Emulex    │
  └──────────────┘   └──────────────┘                 │  Controller  │
                                                      └──────────────┘
  ┌─────────────────────────────────┐
  │      System Industries          │                 ┌──────────────┐
  │    9900 Disk Controller         │                 │Disk Controller│
  └─────────────────────────────────┘                 │  and Drive   │
                                                      │  DEC RK07    │
  ┌──────────────┐   ┌──────────────┐                 └──────────────┘
  │ Fujitsu Eagle│   │ Fujitsu Eagle│
  │ 414 Mbyte    │   │ 414 Mbyte    │                 ┌──────────────┐
  │ Disk Drive   │   │ Disk Drive   │                 │   Ethernet   │
  └──────────────┘   └──────────────┘                 │   UNIBUS     │
                                                      │  Interface   │
                                                      └──────────────┘

                                                      ┌──────────────┐
                                                      │    DZ-11     │
                                                      │ Line Scanner │
                                                      │   8 Lines    │
                                                      └──────────────┘
```
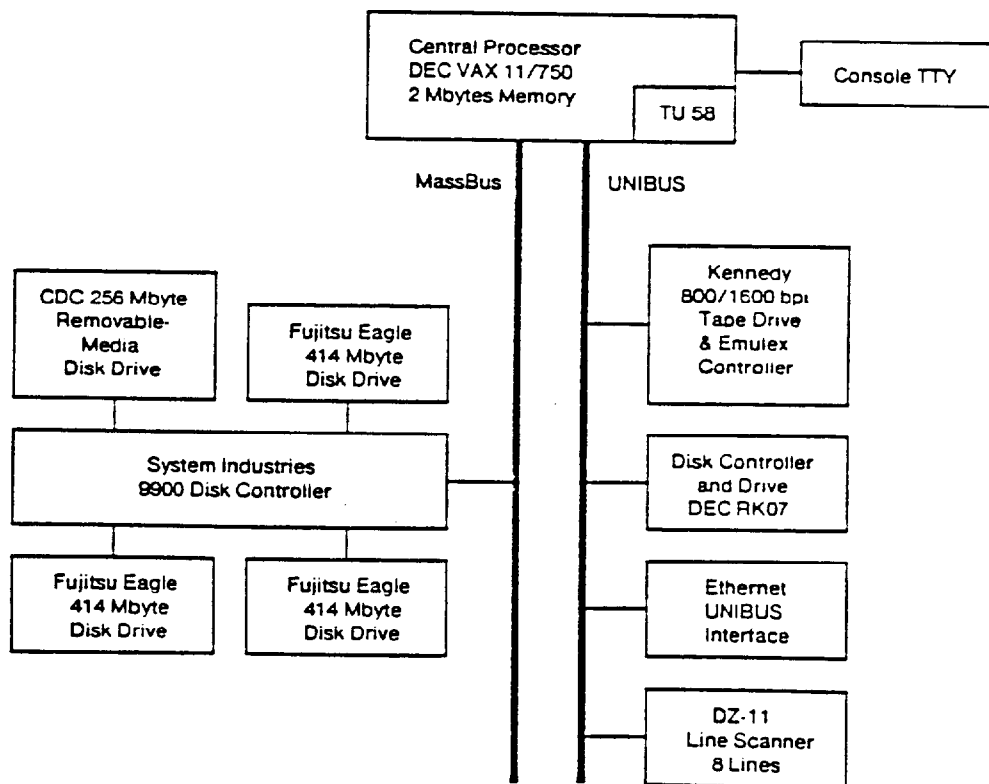
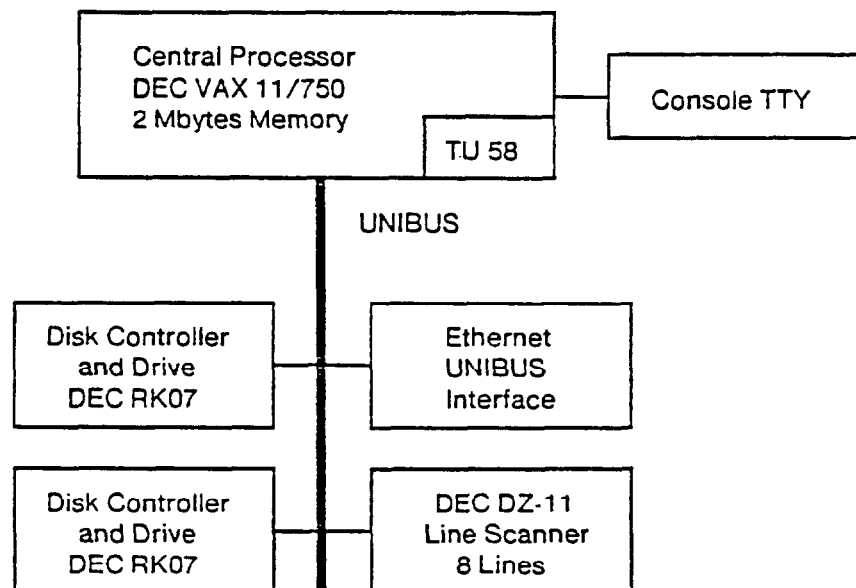**Figure 5:** SUMEX-AIM File Server {SAFE}

Figure 8:   SUMEX-AIM Development Vax {ARDVAX}

I.A.2.3. Timesharing Systems

Continued support and development of our timesharing systems this past year has concentrated on several areas, including improvement of user services such as printing spoolers and archiving support, implementation of features from our KI10 Tenex system, enhancing network interface service, correcting encountered system bugs, and implementing new features for better user community support. In addition, we have invested further effort in supporting the VAX/UNIX system in conjunction with the SUMEX-AIM file server installation.

DECsystem 2060/TOPS-20 System

Support of our main timesharing machine, the DECsystem 2060, has continued during grant year 11.

- *Hardware development*

    1. The DECsystem 2060 system now in operation at SUMEX differs greatly from our previous KI10 Tenex system. Whereas before, the KI10 system and TENEX software required much in-house development and support, life is easier with the 2060. Being at Stanford University, where there are at least 7 other DECsystem 20's with similar hardware and software is a great advantage. We are able to share our experiences with other sites, and have become an integral part of the Stanford DEC community. In addition, the DEC2060 hardware has been more reliable and easier to maintain than the KI10 system.

    2. Additional modems were added to the 2060 to provide support for BELL 212A 1200 baud protocols. This adds an alternative to Vadic 1200 baud service. Modems which use the Bell 212A standard are more widely available for much less cost than Vadic modems.

- *TOPS-20 Monitor Software Enhancements*

    1. A significant enhancement to our TOPS-20 monitor occurred this year when we implemented the software from our Tenex system which allowed extended support for the '?' feature of TOPS-20 when parsing filenames. This feature allows a user at any time to get a list of possible choices when needing to input a file name by just typing '?'. This returns an actual list of file names, whereas in the standard TOPS-20 monitor, just the string 'input filespec' was returned. This is a very significant and useful change to our TOPS-20 system.

    2. We continued to keep up-to-date on the various bug fixes and monitor improvements that we received from DEC and other TOPS-20 sites. These included several fixes and rewrites to the Internet IP/TCP code which went under a major revision this past year.

    3. We installed the capability for users to access their subdirectories as if they were the owners of such. While this may seem to be the logical way to implement subdirectories to begin with, DEC's models of subdirectories was a bit different. Our changes have since been installed on other DEC20's at Stanford and elsewhere.

    4. We installed the capability to vary the allocation of windfall cycles in

accordance with user classes. This will allow us more flexibility in assigning jobs enough cycles to run comfortably while limiting their usage to a strict percentage of the machine.

5. We installed several new features in our TOPS-20 EXEC to facilitate the use of the system by the users. Among these features was the ability to edit any previous command you had entered, and then re-execute that command. This saves extra keystrokes and has proven to be very useful. The code to do this came from the University of Texas at Austin.

6. We switched our system this year to using encrypted passwords. This means that passwords are not stored in any readable form on the computer system, and if an illegal user should gain access to the system, he/she would not be able to find out the passwords of any other users. We feel this feature is quite important as the frequency of computer break-ins/attempts increases.

7. Software was added to our monitor in order to record the last reader of a file. Previously, only the *date* of the last read was recorded, while both the writer and date were recorded for creating and writing a file. This gives users the ability to determine which other persons may have been reading files.

● *Printer Support*

The support on the 2060 for various printers in the SUMEX community has been greatly enhanced this past year.

1. Support was added for the Xerox Raven printer at the Welch Road facility to allow spooling and direct output to the printer. In addition, code was added to the spooler to print out a header page identifying the user, filename, and date for each job.

2. Similar spooler support was added for the Xerox Dover printer in Margaret Jacks Hall.

3. SUMEX installed a Printronix line printer at Welch Road to allow users to print out files remotely from the 2060. The Printronix is connected to SUMEX via a twisted pair serial line.

4. We transformed TENEX software to the normal TOPS-20 line printer spooler program to look out for users who had accidentally printed an 'unprintable file', meaning a binary file of some sort which does not contain legible characters. We do this both by counting the number of binary characters in the first page of the file, and by not printing the file if the count exceeds a certain threshold. A similiar scheme is also used taking into account the vertical motion of the first page.

5. Additional modifications were made to the LP10 line printer driver software in the TOPS-20 monitor to improve the reliability of using this line printer, which came from our KI10 system.

6. We greatly enhanced our support of the IMAGEN Imprint-10 laser printer this past year. A new IP/TCP Ethernet interface was installed on the printer (discussed further in Section I.A.2.4 replacing the existing serial interface. This new interface allows for more efficient printer

operation, and greater flexibility in choosing output modes, number of copies, header pages, and other features. We have also implemented a TOPS-20 spooler for the Imagen as well.

- *User System Software*

    1. We have continued to assemble and maintain a broad range of utilities and user support software on the 2060. These include operational aids, statistical packages, DEC-supplied programs, text editors, text search programs, file space management programs, graphics support, text-formatting and justification assistance, magnetic tape conversion aids, and many more. We also are importing software tools and packages wherever necessary to avoid reinventing the wheel and wasting our own efforts. Packages have been imported from Texas Instruments, Columbia University, the University of Texas at Austin, Yale University, and other Stanford sites.

    2. SUMEX has continued to provide to its users the latest releases of various Lisp dialects that run under TOPS-20. This past year we agreed to provide disk space to store the 'official' version of Interlisp-20 from XEROX due to the fact that the machine at Xerox used to support Interlisp before was being removed. Interlisp-10 is now officially maintained at SUMEX by our staff and XEROX personnel. In addition, we continue to support the full variety of LISPUSERS packages. Portable Standard Lisp (PSL) developed at the University of Utah has also been installed on SUMEX.

    3. We continue to use MM, a very powerful and flexible mail system, on the 2060. Electronic bulletin boards are also extensively-supported at SUMEX. These provide a rather informal mechanism for community discussions and debates. Other bulletin boards, read and contributed to throughout the INTERNET community, are available for perusal at SUMEX. These bulletin boards cover such topics as AI Discussions, Micro Computers, Terminals, and Workstations.

    4. SUMEX participates with other Stanford sites in a general license for access to the SCRIBE text-formatting system from UNILOGIC, including versions to run under TENEX, TOPS-20, and UNIX. SCRIBE is the preferred tool for text preparation at SUMEX.

    5. Versions of various user utilities and system utilities were updated throughout the year. These programs included network server processes, statistical packages, system daemon programs, and several programs for processing electronic mail.

    6. Various system network tables and networking software were updated to accommodate the ARPANET split that occurred this year. The network change effectively split the ARPANET into two networks, the MILNET, which is a secure private part, and the rest of the ARPANET, which operates as the original ARPANET did before.

- *Documentation and Education*

    We have expended considerable effort to develop, maintain, and facilitate access to our documentation so to accurately reflect available software. The

HELP and Bulletin Board subsystems have been important in this effort. As subsystems are updated, we generally publish a bulletin or small document describing the changes. As more and more changes occur, it becomes more difficult for users to track down all of the change pointers. Within manpower limits, we are in a continual process of reviewing the existing documentation system for compatibility with the programs now on line and to integrate changes into the main documents. This also will be done with a view toward developing better tools for maintaining up-to-date documentation.

- *Software Sharing*

   1. As stated previously, we firmly believe in importing rather than reinventing software where possible. As noted above, a number of the packages we have brought up are from outside groups. Many avenues exist for sharing between the system staff, various user projects, other facilities, and vendors. The advent of fast and convenient communication facilities coupling communities of computer facilities has made possible effective intergroup cooperation and decentralized maintenance of software packages.

   2. The TENEX, TOPS-20, and UNIX sites on the ARPANET have been a good model for this kind of exchange based on a functional division of labor and expertise. The other major advantage is that as a by-product of the constant communication about particular software, personal relationships between staff members of the various sites develop. These collegial interactions serve to pass general information about software tools and to encourage the exchange of ideas among the sites. Certain common problems are now regularly discussed on a multi-site level.

   3. We continue to draw significant amounts of system software from other ARPANET sites, reciprocating with our own local developments. Interactions have included mutual backup support, experience with various hardware configurations, experience with new types of computers and operating systems, designs for local networks, operating system enhancements, utility or language software, and user project collaborations. We have been able to import many new pieces of software and improvements to existing ones in this way. Examples of imported software include the message manipulation program MM, SAIL, PASCAL, SOS, INTERLISP, the C compiler, VAX Ethernet code, the PHOTO program, ARPANET host tables, various user utilities, and many others.

   4. Finally, we also have assisted groups that have interacted with SUMEX user projects in acquiring access to software available in our community. We are repeatedly providing tape preparation and copy service to many SUMEX-AIM projects to aid in sharing their software with outside requestors.

DECsystem 2020/TOPS-20 System

1. Monitor Upgrade -- Our 2020 system has continued to run very reliably this past year. We have updated the 2020 monitor with bug fixes and performance improvements regularly. There will likely be few further monitor releases for the 2020 since it does not support extended addressing and there are no plans to add this feature.

2. Demo Controls -- We continue to use the 2020 system for demos of AI systems developed at SUMEX. This demo system takes advantage of the "pie-slice" scheduler in the TOPS-20 release 4 monitor. We now guarantee dedicated users a large fraction of the machine but also allow others to do useful work when the demo demand is low. This system has nicely met the needs of both groups.

VAX/UNIX Systems

We continued to provide systems support for the VAX/UNIX 11/780 system (named 'AIMVAX') shared by the SUMEX-AIM community and Stanford Computer Science Department. Various efforts included supporting the UNIX monitor, installing new network software, and in bringing up various user subsystems.

Further development has continued in support of the SUMEX-AIM File Server (SAFE) based on a VAX 11/750. We successfully converted SAFE to Berkeley Unix 4.2 server, and with the help of the Computer Science Department, converted the Ethernet Pup software to run under UNIX 4.2

I.A.2.4. Professional Workstations

Our ongoing movement to professional workstations is taking on several forms. We continue to carry out our acquisition plans for acquiring Lisp machines for use by the SUMEX community, as well as investigating the use of remote virtual graphics and new lower cost workstations in our environment. This work is prototypical of what other groups will face and we hope will serve to find effective solutions to common problems.

Lisp-based Scientific Workstations

SUMEX carefully developed and implemented our equipment acquisition plan for year eleven by buying seven Xerox 1108 Lisp machines for use by SUMEX-AIM projects. Two of these machines were purchased with special upgrade packages to provide floating point capability, expanded microcode, and expanded memory. Our experiences with these machines will be reported in next years report.

The XEROX Dolphin on loan to Rutgers University was returned to SUMEX this year. This Dolphin had effectively served the Rutgers-AIM community in setting up their Ethernet network and provided initial exposure to the Lisp machine technology. Now that that experiment is successfully completed, the Dolphins will be used for AI system development at Stanford.

SUMEX installed two SYMBOLICS 3600 Lisp machines, purchased with DARPA funding, for use within the Heuristic Programming Project (HPP) at their new location at Welch Road. We are currently awaiting a new release of the Symbolics Operating System software before we can provide Ethernet access to our file server from these machines. The 3600's are used regularly by members of the HPP.

We still are using 4 preproduction models of the Dolphin workstations. One preproduction model has been exchanged for a production system, and we are on schedule with XEROX to exchange the remaining 4 machines for production models at no extra cost. This process is hampered by the rate at which XEROX themselves can get production machines.

We studied the benefits of buying the extended memory and microstore upgrades to the Xerox 1108 Dandelion announced at AAAI-83 as being "under development." We concluded that some users would benefit greatly from these enhancements and others not at all. The most marked improvements came from system which were extremely memory limited, such as NEOMYCIN. SUMEX will be acquiring two 1108's with the upgrades for full time use and testing.

A close relationship between SUMEX and the newly-formed Center for the Study of Language and Information (CSLI) at Stanford was established. This has already benefited SUMEX (and the ONCOCIN project in particular) in the loan, by CSLI to SUMEX, of two Xerox 1108's which have been in constant use by researchers since January 12th. The SUMEX staff assisted CSLI in bringing their DecSystem20 and network environment on-line. CSLI has informally expressed an interest in working on the problem of distributed AI computation with SUMEX researchers. CSLI will have 110 1108's on the Ethernet within the year. This resource suggests some exciting solutions to former compute-bound problems. The ONCOCIN group has already implemented a preliminary network-based Interactor which permits elements of ONCOCIN to run concurrently on different machines. As of this writing, the Reasoner and its Debugger have been made to run transparently in this mode, and to make good use of both processors.

E. A. Feigenbaum

Virtual Graphics

SUMEX continued the development of a Virtual Graphics system written in Interlisp-10 and running on our 2060. Any user running the V system on a workstation can then use the package on the 2060 to drive the graphics display on the workstation. A current application is to take nuclear magnetic resonance data on the 2060 and display the atoms and their bonds on a SUN workstation by using splines. This development is in its infancy, but is opened ended and has great potential with the price of workstations capable of decent graphics reaching the two to three thousand dollar range. It allows those users who cannot afford expensive lisp machines to have full graphics power available to them by doing the actual graphics applications on a large time shared system, and then doing the graphics itself remotely on a less expensive workstation. This development can help users take advantage of the computing power of the DECsystem 20, while providing many of the high speed graphics advantages of the Lisp Machines.

Apple Workstation Development

SUMEX-AIM has initiated a development project to pursue the effectiveness and possible use of low cost personal workstations within our environment. After examing a number of new personal computers and workstations on the market, such as the Hewlett Packard 150, IBM PC, Sun workstations, and others, we chose the Apple Macintosh and Apple Lisa on which to begin our work in this area. These machines were chosen technically due to their built in graphics, networking, mouse, windows, and menu support. We also considered the very beneficial relationship formed between Apple and Stanford University which provides us direct access to Macintosh hardware and software documentation which is a necessity for the type of work we plan to do.

Our Macintosh development encompasses several areas ;

1. *INFO-MAC Discussion List*

   An electronic discussion list was originated at SUMEX, and is currently maintained here, to foster sharing and communication among research groups and universities that are interested is pursuing the serious use of the Macintosh within their respective environments. This list has been highly successful in collaborating on Macintosh development and the sharing of ideas. The discussion list currently contains over 50 sites, and well over 1000 participants.

2. *C Development Environment*

   A vital link in our development of Macintosh software is creating a C based development system on our VAX computers for the coding and downloading of software. Utilizing existing MC68000 C cross compilers on our VAX, we are developing the necessary linkages in order to make the appropriate system calls to routines in the Macintosh ROM's for sophisticated graphics and system related functions.

3. *Macintosh print servers*

   In order to effectively use the Macintosh as a stand-alone workstation, we will provide the ability to print out Macintosh developed files on our IMAGEN laser printers.

4. *Applebus to Ethernet Interface*

   This development involves the hardware and software necessary to be able to

access various file and print servers on our Ethernet from a Macintosh. The Macintosh will be connected to the Apple network called Applebus. Our hardware provides an interface between Applebus and 10MB Ethernet. The software necessary for this project involves formulating Macintosh file level and block level I/O requests into properly formatted Internet packets.

5. *Virtual Graphics on a Lisa*

The Virtual Graphics System, as previously reported, is in great need of a low cost workstation on which it can run. We have started a project to port the Virtual Graphics system to a Apple Lisa in hopes of providing to our users high speed graphics at remote locations. We will report further on this project in next year's report.

Anticipating the popularity of our Macintosh developments, we are fully prepared to make our efforts available to other research sites, Universities, and non-profit institutions on a royalty-free basis in hopes of fostering continued development and communal sharing.

In addition to Lisp-based scientific workstations, we believe the use of low cost workstations, which offer suitable local processing power, high resolution screens with easy to use user interfaces, and networking and communications abilities, are vital to the future of our resource. Our Macintosh and Lisa development efforts will allow us to use and experiment with these workstations in our environment.

Hewlett Packard Development

SUMEX assists the ONCOCIN Project is developing a computing environment for developing AI applications based on HP 9836 workstations. These workstations were part of a gift from Hewlett Packard to the Oncocin Project. Additional support peripherals for the 9836's included large capacity disk drives, color monitors, graphic tablets, and a laser printer. Work is proceeding to network these machines onto the SUMEX Ethernet as soon as suitablke networking hardware is available from HP. These machines will be used for new and existing projects within Oncocin.

I.A.2.5. Networking and Communications

A highly-important aspect of SUMEX-AIM is effective communication with remote users and between the growing number of machines available within the SUMEX resource. In addition to the economic arguments for terminal access, networking offers other advantages for shared computing, including improved inter-user communications and more effective software sharing.

Users accessing a remote computer will use a hardline connection to the computer as a standard of comparison. Local networks stand up well in this comparison but remote network facilities do not. Data loss is not a problem in most network communications; in fact, with the more extensive error checking schemes, data integrity is higher than for a long distance phone link. On the other hand, remote networking relies upon shared use of communication lines for widespread geographical coverage at substantially reduced cost. However, unless enough total line capacity is provided to meet peak loads, substantial queueing and traffic jams result in the loss of terminal responsiveness. We continually monitor the load statistics for our direct, dialup, and TYMNET lines to avert logjam situations.

TYMNET

TYMNET provides broad geographic coverage for terminal access to SUMEX from throughout the country and increasingly from foreign countries. With the installation of our new DEC2060 computer system in January of 1983, we installed new TYMNET equipment. After the initial debugging of the new equipment (called TYMCOM) the equipment has been quite reliable. However, some months after the installation it was discovered that the XON/XOFF protocol between the Tymcom and the 2060 had not been properly specified in the Tymcom and was corrected. The number of user complaints about connection problems have been greatly reduced. This is thought to be the result of improved "backbone" lines within Tymnet and the installation of triple-duty modems which simplify things for the users.

ARPANET

We retain our advantageous connection to the Department of Defense's ARPANET, now managed by the Defense Communications Agency (DCA). This connection has facilitated close collaboration with the Rutgers-AIM facility and many other computer science groups that are also on the net. We have maintained good working relationships with other sites on the ARPANET for system backup and software interchange. Such day-to-day working interactions with remote facilities would not be possible without the integrated file transfer, communication, and terminal-handling capabilities unique to the ARPANET. The ARPANET is also key to maintaining on-going intellectual contacts between SUMEX projects such as the Stanford Heuristic Programming Project authorized to use the net and other active AI research groups in the ARPANET community.

This past year, SUMEX-AIM participated in the split of the ARPANET into two networks; the MILNET, which is a highly secure strictly DOD-related part of the network, and the ARPANET, which is the remainder of the ARPANET sites. This latter net functions as we knew the ARPANET before. The MILNET can only be accessed via mail gateways. No TELNET or FTP to MILNET sites is allowed. In addition, access to the ARPANET TAC's (Terminal Access Controllers) was restricted this past year to only those users who were granted TAC access cards, which meant their username was registered with the Network Information Center, and they were given a password with

which they could dial into the ARPANET. SUMEX arranged for guest cards for those users who needed such access.

We continue to be called upon to interact with outside organizations which are (or wish to be) connected to our IMP. The line to Advanced Information and Decision Systems occasionally causes trouble requiring diagnosis. The intended connection to Perceptronics Inc. has evidently been canceled.

ETHERNET

A substantial portion of our system effort this past year went into continued development of local Ethernet facilities which link the SUMEX resource hardware with other parts of the campus, namely to 701 Welch Road, which is the new location of the Heuristic Programming Project, and to the Computer Science Department building on campus. We have also invested a great amount of effort this year to begin our transfer to a 10 megabit Ethernet, while continuing support of our current 3 MB ethernet.

Specific areas of Ethernet development include:

1. *Leaf server* -- We continued support of the Sequin reliable packet protocol and Leaf byte-level file transfer protocol to enable our Xerox D machines to access files on our DEC20 systems. The Leaf server had to be modified on the DEC20 this past year when we switched to using encrypted passwords. The LEAF server implementation for the 4.2 BSD release of Unix was also debugged and installed at SUMEX. This allows us to access files stored on our VAX file servers from either our 10MB or 3MB networks.

   The Leaf protocol is built into the lowest levels of the Dolphin I/O system, and allows any file on a remote file server to be accessed as easily as a disk file in both paged or random access mode. The latest updates to the Sequin transport level have made marked improvements in efficiency. The 2020 now performs Leaf file transfers with a speed approaching that of XEROX's dedicated file server.

2. *TOPS-20 Ethernet Server* -- We continued to maintain and improve the Ethernet service under the TOPS-20 operating system. This included updates to the TELNET and FTP programs, as well as mail software, the previously mentioned Leaf server, and network table maintanence programs.

3. *Ethernet Gateway* -- Our Ethernet gateway software has continued to run reliably and effectively. The previous problems with lost packets and delayed terminal response has been fixed, the cause of which was a bad memory board and a software bug in the TOPS-20 operating system. Serious problems that affected our net connectivity to other parts of campus were also discovered and repaired this past year, thereby providing us with over 99% net connectivity to the rest of campus. The changes involved board repair and modifications to the topology of the campus Ethernet.

   The gateway itself was generalized to handle 3 or more directly connected networks where previously it had only dealt with 2 such networks. We currently have two gateways, each handling the traffic between three local networks, two of which are 3MB and one a 10MB network.

4. *Ethernet TIP (EtherTIP)* -- The EtherTIP provides multiple terminal access to the Ethernet. A PUP ethernet operating system was written for MC68000-based processors, and a MC68000-based EtherTIP was built based on this.